Topological Fuzzy Pharmacophore Triplets and adapted Molecular Similarity Scoring Schemes

Dragos Horvath

UMR 8576 CNRS - Université des Sciences et Technologies de Lille

Bât. C9 - 59655 Villeneuve d'Ascq CEDEX, France

This work introduces a novel molecular description – topological (2D) Fuzzy Pharmacophore Triplets, 2D-FPT, using the number of interposed bonds as the measure of separation between the atoms representing pharmacophore types (hydrophobic, aromatic, hydrogen bond donor and acceptor, cation, anion). 2D-FPT feature three key improvements with respect to the state-of-the-art pharmacophore fingerprints:

- Fuzzy mapping of molecular triplets onto the basis set of pharmacophore triplets: Unlike in the binary scheme where an atom triplet is set to highlight the bit of a single, best matching basis triplet, the herein defined fuzzy approach allows for gradual mapping of each atom triplet onto several related basis triplets, thus minimizing binary classification artifacts.
- Proteolytic equilibrium-dependence, by explicitly considering all the conjugated acids and bases (microspecies). 2D-FPT are concentration-weighted (as predicted at pH=7.4) averages of microspecies fingerprints. Therefore, small structural modifications, not affecting the overall pharmacophore pattern (in the sense of classical rule-based assignment), but nevertheless triggering a pK_a shift, will have a major impact on 2D-FPT. Pairs of almost identical compounds with significantly differing activities ("activity cliffs" in classical descriptor spaces) were in many cases predictable by 2D-FPT.
- A new similarity scoring formula, acknowledging that simultaneous absence of a triplet in two molecules is a less constraining indicator of similarity than its simultaneous presence. It displays excellent Neighborhood Behavior (NB), outperforming 2D or 3D two-point pharmacophore descriptors or chemical fingerprints.